

MANORAA (Mapping Analogous Nuclei Onto Residue And Affinity) for identifying protein–ligand fragment interaction, pathways and SNPs

Duangrudee Tanramluk^{1,2,*}, Lalita Narupiyakul^{2,3}, Ruj Akavipat^{2,4}, Sungsam Gong⁵ and Varodom Charoensawan^{2,6}

¹Institute of Molecular Biosciences, Mahidol University, Salaya, Nakhon Pathom 73170, Thailand, ²Integrative Computational BioScience (ICBS) Center, Mahidol University, Salaya, Nakhon Pathom 73170, Thailand, ³Department of Computer Engineering, Faculty of Engineering, Mahidol University, Salaya, Nakhon Pathom 73170, Thailand, ⁴Department of Computer Science, Faculty of Science, Kasetsart University, Chatuchak, Bangkok 10900, Thailand, ⁵Department of Obstetrics and Gynaecology, University of Cambridge, The Rosie Hospital, Cambridge CB2 0SW, UK and ⁶Department of Biochemistry, Faculty of Science, Mahidol University, Ratchathewi, Bangkok 10400, Thailand

Received February 8, 2016; Revised April 7, 2016; Accepted April 13, 2016

ABSTRACT

Protein–ligand interaction analysis is an important step of drug design and protein engineering in order to predict the binding affinity and selectivity between ligands to the target proteins. To date, there are more than 100 000 structures available in the Protein Data Bank (PDB), of which ~30% are protein–ligand (MW below 1000 Da) complexes. We have developed the integrative web server MANORAA (Mapping Analogous Nuclei Onto Residue And Affinity) with the aim of providing a user-friendly web interface to assist structural study and design of protein–ligand interactions. In brief, the server allows the users to input the chemical fragments and present all the unique molecular interactions to the target proteins with available three-dimensional structures in the PDB. The users can also link the ligands of interest to assess possible off-target proteins, human variants and pathway information using our all-in-one integrated tools. Taken together, we envisage that the server will facilitate and improve the study of protein–ligand interactions by allowing observation and comparison of ligand interactions with multiple proteins at the same time. (<http://manoraa.org>).

INTRODUCTION

Understanding protein–ligand interaction is crucial for drug discovery research, as it defines the binding affinity, steric complementarity of the surface and pharmacophoric patterns of the compound to the target protein. Favorable ligand interactions with protein such as suitable polar

groups counterparts and proper hydrogen bonding partners are crucial for the ligand design process and the imperfect fit between the protein and the ligand will result in decreased binding affinity (1). A number of tools is available for visualizing and analyzing protein–ligand interaction; however, only few can provide comprehensive information such as verified binding affinity, and couple the results with the ligand interaction visualization available for multiple protein comparison in the same place (2). By understanding the favorable interactions between the target protein and a ligand of interest, one can start to rationalize drug design strategy and make the protein engineering possible by strengthening preferred interactions for instance.

To date, there are more than 100 000 structures in the Protein Data Bank (PDB) (3). However, it is not always straightforward to harness all the relevant information from the PDB. Querying the substructure of the ligands to return multiple molecular interactions that are available in the PDB can take a considerable amount of time as one normally goes through a series of non-intuitive steps. After multiple protein–ligand structures are retrieved, the comparison can be complicated and time-consuming, especially when the structures contain a large amount of protein–ligand interactions from multiple contact points, which normally have to be investigated individually and manually. To the best of our knowledge, there is no existing tool specifically designed for comparative analysis of protein–ligand interactions in multiple structures at the time.

Two of the most popular tools for searching molecular interactions in the binding sites are Relibase (4) and PDBe-Motif (5). Both tools are restricted to the structures in the PDB and are often used to show the distribution of protein–ligand binding patterns in the PDB as a whole. Other tools such as PLIP (6) are also available for investigating protein

*To whom correspondence should be addressed. Tel: +66 2 441 9003 (Ext. 1211); Fax: +66 2 9428499; Email: duangrudee.tan@mahidol.ac.th

ligand interactions and visualization; however, the users cannot obtain knowledge of preferred interactions easily because it is dedicated to visualization and does not allow sorting by binding affinities, or viewing multiple proteins structures that bind to the same ligand in the same panel. The PLIC database (7) provides protein–ligand interaction clusters and also other related binding site information, and also has a superposition panel based on the clustering of similar binding sites. However, the ligand superposition is performed as a whole molecule, not based on the equivalent substructures, and hence, it is difficult to directly relate that information to the change in the binding affinities. WONKA (8) on the other hand, can offer observation from multiple structures but it requires the users to supply the set of superposed proteins with their equivalent amino acids re-named to the corresponding residue numbers. PoSSuM (9) aims to detect similar small molecule binding pockets; however, the overall similarity between pockets do not guarantee the same ligand binding pattern. A tool such as PLI (10) can also be used to find a particular ligand binding to a list of homologous proteins.

A direct query of ligand to the RCSB Protein Data Bank (3,11) returns the data retrieval in the form of PDB files and Jmol applet but does not provide ligand substructure analyses for multiple structures. The databases BindingDB (12) and Binding MOAD (13) emphasize the binding affinities data for further use such as for QSAR analysis (14), which does not offer the structural analysis of the binding site or the trend of binding affinity. In addition, these databases do not provide links from fragments to pathways or known human variants such as SNPs, a feature that will be useful for the drug design in the personalized medicine era.

To this end, we have developed the integrative web server MANORAA (Mapping Analogous Nuclei Onto Residue And Affinity) to facilitate understanding of ligand selectivity and promiscuity through the analysis of multiple protein structures on the web interface. It enables researchers to retrieve multiple chemical compounds and their binding partner proteins from the PDB, and compare and visualize the ligand-residue contact interactions all at the same time. Other useful functionalities include sorting of binding affinities of multiple proteins, as well as obtaining additional information such as protein functions, the species that a particular ligand is found in a complex, and the pathways that the ligand is found to take part in by linking to a pathway map such as KEGG (15), all in one place.

MANORAA: rationales, input and output

We built MANORAA with an aim to provide a user-friendly, one-stop service for ligand–protein interaction investigation. MANORAA was developed on top of CREDO, a database devoted specially to the protein–ligand interaction, which provides all pairwise atomic interaction contacts between ligand and proteins from the PDB in the form of a relational database (16,17). By filtering and ranking the interaction types in a systematic manner, the ligand contacts that are most important can be shown and can be related to the change in the binding affinities. This server provides integrated information about the target and off-target proteins interacting with the query ligand from the

latest publicly available mirror of CREDO. MANORAA also provides protein–ligand binding affinity values from the Binding MOAD database (13) where the high quality binding affinity data are collected from literature. Importantly, the users can observe and compare the interaction by processing the ligand contact with multiple protein structures based on the complexes deposited to the PDB all at the same time. All queries to MANORAA start from one simple input page, and the results are provided in two output steps, as described here. We have extensively tested MANORAA on several common operating systems and web browsers (see Supplementary Data for details), and the most compatible browsers that we recommend are Windows version 7 or higher, and OSX Maverick or higher, and on Chrome 49 or higher, and Safari 9 or higher.

Input: chemical structure

The users can start with a ligand or part of a ligand of interest by providing one of the following as an input: (i) chemical name, (ii) SMILES expression, (iii) PDB ligand's 3-letter code or (iv) chemical structure (Figure 1). To facilitate generation of a SMILES expression, the MANORAA provides the SMILES lookup and then exports the SMILES to the chemical sketch panel as shown in Figure 1. The users can also select to create or edit a SMILES expression by drawing a chemical structure, or modify some parts and then import that sketch to a SMILES string before submission. From SMILES, the users can link to extended ligand names and compound bioactivity information via ChEMBL (18). The web server employs a JavaScript library called MarvinJS from ChemAxon to achieve the task. MarvinJS provides an HTML5-based user interface for chemical drawing, which allows the users to have an interactive interface without the need to install any additional plug-in.

Output 1: list of proteins interacting with a queried ligand

Once the users submit a ligand, a list of PDB entries that contain the submitted ligand will be returned. Ligand(s) of similar chemical structure and their target proteins will be returned if any part of the molecule matches with the SMILES input fragment. For example, Figure 2 shows a table of PDB entries interacting with a ligand 'STU' (Staurosporine). The binding affinity values, taken from Binding MOAD (13), are provided to help prioritize target proteins as they imply the binding strengths between the query ligand and the targets. For each entry, the following external information is also provided: (i) the pathway information from KEGG (15), (ii) the protein information from UniProt (19), (iii) the amino acid variants from SAMUL (20) and (iv) variants, isoforms and genomic context, protein/RNA baseline expression, gene ontology from the Centre for Therapeutic Target Validation (Open Targets, <https://www.targetvalidation.org/>). The server links crystallographic structures of protein–ligand interaction to related biochemical pathways via UniProt ID to KEGG ID mapping. The user can also link the proteins of interest to known human variants such as SNPs in the coding regions via the SAMUL web server (20). This essentially allows researchers to predict whether the candidate ligands will have a tendency to bind proteins with different annotated SNPs in the

The screenshot shows the 'Query Method' section of the web server. It includes a search bar for chemical names, radio buttons for 'SMILES' and 'PDB Ligand', a SMILES input field, a PDB ligand code dropdown, and a 2D chemical structure drawing tool. Arrows from external labels point to these specific input areas.

Figure 1. Input panel for ligand fragment via chemical names, SMILES expression, PDB ligand's three-letter code or chemical structure drawing.

coding regions, an important step of drug design in the personalized medicine era. These results can be exported as a CSV table.

This result page serves as an input form for the next step, which is to visualize three-dimensional (3D) structures of protein–ligand interactions based on the selection of (i) atom of interest in the ligand, and (ii) the target PDB chains. To aid this process, the MANORAA web server shows the ligand chemical structure so that users can pick up atoms of their interest interactively. By default, all the heteroatoms of ligands and PDB chains where the binding affinity are available are pre-selected.

Output 2: visualization of ligand–protein interactions

Once the users select ligand atoms and associated protein chains of interest, MANORAA will connect to CREDO to obtain interacting partner proteins of each ligand and PDB pair, grouped by nine interaction types and highlight them by different colors. The available interaction types are aromatic, hydrogen bond, ionic interaction, covalent bond, metal complex, carbonyl interaction, halogen bond, hydrophobic interaction and van der Waals clash (as shown in Figure 3 with criteria in Supplementary Data). The server will then rank the most important contact based on the shortest distances of unique interaction found for every atom per amino acid residue. JSmol (21), a JavaScript framework based on HTML5 for displaying interactive 3D molecular structures, has been employed into the user inter-

face to enable the users to toggle display of the interaction partners in the 3D viewing panel. Display of ligand–protein interaction at each residue can be obtained by clicking the loading button of each PDB IDs, then choosing the residue name of interest. This step allows the user to have full control on what part of the chemical structure that they want to focus on. The results can be revisited using a unique URL provided. The list of target proteins and contact residues can be printed as PDF file together with the protein structures which can be saved from JSmol. Note that additional technical details of the web server can be found as Supplementary Data.

To assist the first-time users, we have provided comprehensive step-by-step tutorial, demo video and sample pages on the web server. Here, we also provide two examples of how MANORAA can be employed to assist real-world drug design and protein engineering research.

Making use of MANORAA in ligand–protein interaction studies

Case study 1: the trend of interaction observed in N4 of STU interacting with the kinase family. To illustrate the use of MANORAA and its features, here we use our previous comprehensive study on staurosporine's binding strength as an example (22). The study demonstrated that staurosporine's strength of interaction with kinase depends on the number and the orientation of hydrogen bonds and ionic interactions made around the N4 atom of staurosporine.

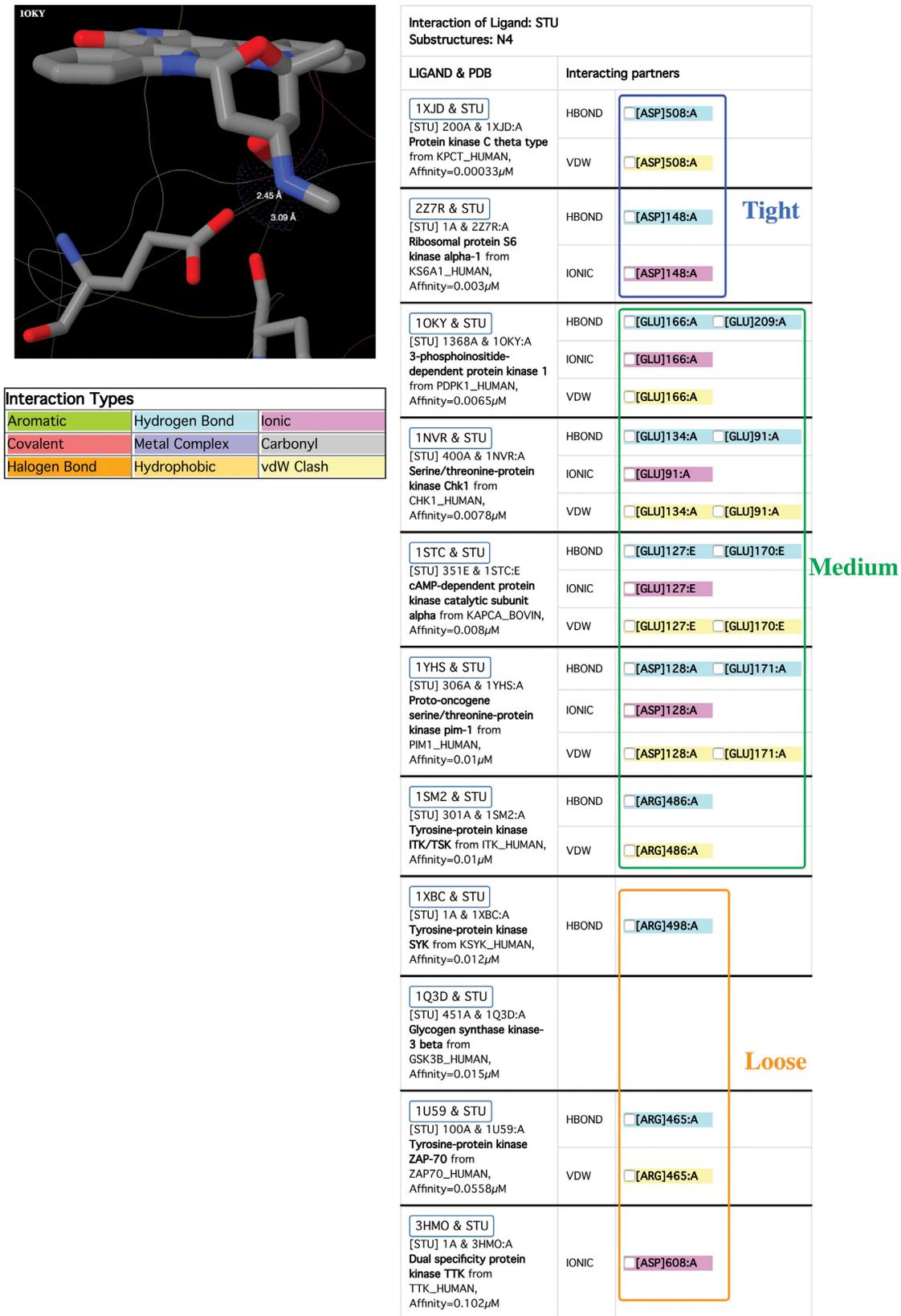


Figure 3. Display of the ligand three-letter code STU (staurosporine) interacting with several proteins in the kinase family. The top row shows interaction with the lowest binding affinity value, which means staurosporine can tightly bind (0.00033 µM) via aspartate, and there are at most two hydrogen bonds (cyan) or ionic interaction (magenta) in combination. The second group has binding affinity values between 0.0065 to 0.010 µM and has three interactions that are either ionic or hydrogen bonds. All the others with binding affinity value more than 0.010 µM have one hydrogen bond, or ionic interaction, or no interaction at all.

For the worst STU binding to kinase cases, N4 from PDB IDs: 1XBC, 1U59, 3HMO have only one bond in the non-preferable orientation while the one from 1Q3D does not have any hydrogen bonding or ionic interaction at all. These two structural observations imply poor binding affinity and hence they are present in the group of large binding affinity values. In contrast, 1OKY, 1NVR, 1STC and 1YHS have better binding affinities due to better hydrogen bonding and ionic interaction in both the structural orientation sense and also the amount of bonds they made which are at least two hydrogen bonds plus one ionic interaction. Note that binding via aspartate makes the interaction tighter than binding via arginine as seen in 1XJD and 2Z7R and hence shown with the lowest K_i or tightest binding affinities from the Binding MOAD database. These types of analysis will be beneficial to drug design because we know which part of the ligand is the major determinant of the binding affinities and what amino acids facilitate those preferred interactions. With MANORAA, these processes can be performed all in one place (Figure 3).

Case study 2: how MANORAA can be powerful for analyzing ligand and its binding protein with known SNPs that can be linked to diseases. Trifluoperazine (PDB ligand code TFP) was originally identified as an antipsychotic drug used in the treatment of schizophrenia, via blocking D2 dopaminergic receptors in the brain. However, trifluoperazine has recently been repurposed to inhibit the growth of cancer stem cells via its function as a calmodulin inhibitor, but the inhibitory mechanism is unclear (23,24). Using MANORAA, we can demonstrate that this ligand can also bind with many target proteins such as the placental calcium binding protein (S100-A4), calmodulin and troponin C (in a way, ‘off-targets’ to dopaminergic receptors). The result shows that the N3 atom of TFP interact with the placental calcium binding protein (S100-A4) via one hydrogen bond, or one optional ionic interaction in the cases of bovine calmodulin and human troponin C, suggesting the importance of this atom for all the proteins that bind to this small molecule. On the other hand, the N2 atom of TFP forms two ionic interactions with both bovine calmodulin structures but does not make any significant interaction for either S100-A4 protein or human troponin C. This kind of information can be useful for designing the selectivity of the drug.

Furthermore, MANORAA (via SAMUL) also reveals two SNPs in the human calmodulin gene that have been associated with ventricular tachycardia, a common side effect from trifluoperazine use. MANORAA also provides a list of multiple bovine calmodulin with crystallographic structures that harbor the ligand, enabling researchers to explore the effect of amino acid changes to affect ligand–protein interaction. This demonstrates how MANORAA can be used for an initial assessment of drug repurposing results. It should be noted that our data relies on crystallographic structures deposited to the PDB at the time. Saying that, MANORAA provides another way to make use of the growing PDB by linking the structures to human genome variations.

DISCUSSION AND FUTURE DIRECTIONS

Thanks to technological advances in crystallography and other methods for determining structures of biological molecules, the bottleneck of structural biology is now shifting from obtaining the structures to interpreting and linking them to other biological information such as pathways and genomic variants. With the wealth of information on ligand–protein interactions from publicly available databases such as PDB, it is now possible to perform a comparative study of multiple ligands and proteins (or drug candidate compounds and target proteins) at the same time. MANORAA has been established to facilitate these processes all in one place.

The web server has a number of useful features that assist the investigation of ligand–protein binding specificity, biological pathways the proteins are involved in and known human variants in the coding regions of the proteins. As depicted by Böhm, ligand with poor binding affinity is caused by missing crucial active site interactions in comparison with other tight binding ligands (1). Our service allows the user to compare the ligand’s binding affinities with the numbers and the types of interactions that the ligand makes with multiple proteins, which should be useful for users to identify the key residues of proteins and the atoms of ligands in order to manipulate the interaction strength.

Existing protein–ligand contact and interaction databases are required in order to expedite the process of calculating and classifying the molecular interaction on the fly. To this end, we make use of the CREDO backend, and provide links to SAMUL (20), UniProt (19), PDBe (5), PDBsum (25) and KEGG (15). Note that the total size of calculated interaction for CREDO databases alone, including all the structures in the PDB, is very large (72 GB), but that would allow the interactions to be observed almost instantaneously. The graphical representation on JSmol (21) allows the users to view multiple structures with the fragment in the same window. The color highlight of the protein–ligand interaction that links to the JSmol structure visualization panel in real-time allows robust ligand interaction identification so that the researchers can relate the knowledge of the binding affinity value to the missing or occurring interaction by themselves. MANORAA employs a support responsive design, which means the output structures can be visualized without distortion even on a tablet or mobile phone. Another unique feature of MANORAA is its multiple-structures visualization panel with multiple loading buttons. The users can observe multiple structures one at a time and progress to each one to get an impression of the whole set of proteins that interact with this particular ligand fragment, and could identify the amino acid residue or atoms of a ligand that can be modified to fine-tune the ligand–protein binding interaction.

The main strengths of MANORAA over other previously available web servers aforementioned includes its flexibility of analyzing multiple experimentally verified ligand–protein interactions at the time, using its user-friendly and fast responsive interface. PLIP (6), for instance, focuses on the visualization of one structure at a time, while PoSSuM (9) provides the superposition and makes a comparison be-

tween residues surrounding a protein pair, rather than giving details of the type of residues in contact. Even though there are a number of tools that allow multiple structure observation, the interactions are not dissected to different chemical interaction types and do not provide visualization panel for the users to drill down to the level of ligand substructure interactions.

Looking ahead, we aim to routinely maintain the server and add new functionalities, which will be managed by a programmer dedicated to MANORAA's development and a multidisciplinary team. For instance, we have been developing a new algorithm to show gradient color of atomic position conservation. This will allow us to show position-specific interaction by highlighting the active site based on the percent conservation of the atomic position surrounding the ligand substructure. For the time being, we have implemented this for 18 staurosporine superposed complexes as an example from our sample page (see Supplementary Data). In addition, protein ligand contacts will be updated for every newer release of CREDO. In addition, PDBe (5), PDBsum (25), CACTVS (26,27), ChEMBL (18), KEGG (15), Open Targets (<https://www.targetvalidation.org/>), UniProt (19) and SAMUL (20) are accessed in real-time through their websites, hence the results shown will always be the most updated.

With MANORAA, the chemical fragments that have an influence on different pathways in different organisms can open the door for a more robust and insightful analysis for the study of multi-target drug design, species selectivity, off-target inhibition causing drug side effect problems. We envisage that MANORAA will provide a missing link between structural biology, systems biology and genetics information by one central concept surrounding the ligand's chemical structure to assist drug discovery and the probe molecules community.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We are grateful to Sir Tom Blundell, Dr Adrian Schreyer and Dr Bernardo Ochoa who provided the CREDO database and M.R. Jisnuson Svasti who supported the initial phase of the MANORAA server. We thank Pisitsak Sriwimol and Sakao Phoochaijaroen for their technical supports. We also thank Dr Bhoom Suktitipat and Lect. Stephen Smith for useful comments on the manuscript. Marvin JS 15.2.23, 2015, ChemAxon (<http://www.chemaxon.com>), was used for drawing and displaying chemical structures and substructures.

FUNDING

Mahidol University; Thailand Research Fund via Institute of Molecular Biosciences and Faculty of Science, Mahidol University [MRG5480208 to D.T., TRG5880067 to V.C.]; Crown Property Bureau Foundation via the Integrative Computational BioScience (ICBS) center, Mahidol University. Funding for open access charge: Thailand Research Fund [IRG5780009]; Mahidol University Research Grant.

Conflict of interest statement. None declared.

REFERENCES

- Böhm, H.-J. (2003) *Protein–Ligand Interactions—From Molecular Recognition to Drug Design*. WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.
- Inhester, T. and Rarey, M. (2014) Protein–ligand interaction databases: advanced tools to mine activity data and interactions on a structural level. *Wiley Interdiscip. Rev. Comput. Mol. Sci.*, **4**, 562–575.
- Rose, P.W., Prlic, A., Bi, C., Bluhm, W.F., Christie, C.H., Dutta, S., Green, R.K., Goodsell, D.S., Westbrook, J.D., Woo, J. *et al.* (2015) The RCSB Protein Data Bank: views of structural biology for basic and applied research and education. *Nucleic Acids Res.*, **43**, D345–D356.
- Hendlich, M., Bergner, A., Gunther, J. and Klebe, G. (2003) Relibase: design and development of a database for comprehensive analysis of protein–ligand interactions. *J. Mol. Biol.*, **326**, 607–620.
- Velankar, S., Alhroub, Y., Best, C., Caboche, S., Conroy, M.J., Dana, J.M., Fernandez Montecelo, M.A., van Ginkel, G., Golovin, A., Gore, S.P. *et al.* (2011) PDBe: Protein Data Bank in Europe. *Nucleic Acids Res.*, **40**, D445–D452.
- Salentin, S., Schreiber, S., Haupt, V.J., Adasme, M.F. and Schroeder, M. (2015) PLIP: fully automated protein–ligand interaction profiler. *Nucleic Acids Res.*, **43**, W443–W447.
- Anand, P., Nagarajan, D., Mukherjee, S. and Chandra, N. (2014) PLIC: protein–ligand interaction clusters. *Database*, bau029.
- Bradley, A.R., Wall, I.D., von Delft, F., Green, D.V., Deane, C.M. and Marsden, B.D. (2015) WONKA: objective novel complex analysis for ensembles of protein–ligand structures. *J. Comput. Aided Mol. Des.*, **29**, 963–973.
- Ito, J.-i., Ikeda, K., Yamada, K., Mizuguchi, K. and Tomii, K. (2014) PoSSuM v.2.0: data update and a new function for investigating ligand analogs and target proteins of small-molecule drugs. *Nucleic Acids Res.*, **43**, D392–D398.
- Gallina, A.M., Bisignano, P., Bergamino, M. and Bordo, D. (2013) PLI: a web-based tool for the comparison of protein–ligand interactions observed on PDB structures. *Bioinformatics*, **29**, 395–397.
- Rose, P.W., Bi, C., Bluhm, W.F., Christie, C.H., Dimitropoulos, D., Dutta, S., Green, R.K., Goodsell, D.S., Prlic, A., Quesada, M. *et al.* (2013) The RCSB Protein Data Bank: new resources for research and education. *Nucleic Acids Res.*, **41**, D475–D482.
- Liu, T., Lin, Y., Wen, X., Jorissen, R.N. and Gilson, M.K. (2007) BindingDB: a web-accessible database of experimentally determined protein–ligand binding affinities. *Nucleic Acids Res.*, **35**, D198–D201.
- Benson, M.L., Smith, R.D., Khazanov, N.A., Dimcheff, B., Beaver, J., Dresslar, P., Nerothin, J. and Carlson, H.A. (2008) Binding MOAD, a high-quality protein–ligand database. *Nucleic Acids Res.*, **36**, D674–D678.
- Hansch, C. and Silipo, C. (1974) Quantitative structure–activity relation of reversible dihydrofolate reductase inhibitors. Diaminotriazines. *J. Med. Chem.*, **17**, 661–667.
- Kanehisa, M. and Goto, S. (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.*, **28**, 27–30.
- Schreyer, A. and Blundell, T. (2009) CREDO: a protein–ligand interaction database for drug discovery. *Chem. Biol. Drug Des.*, **73**, 157–167.
- Schreyer, A.M. and Blundell, T.L. (2013) CREDO: a structural interactomics database for drug discovery. *Database*, bat049.
- Bento, A.P., Gaulton, A., Hersey, A., Bellis, L.J., Chambers, J., Davies, M., Kruger, F.A., Light, Y., Mak, L., McGlinchey, S. *et al.* (2014) The ChEMBL bioactivity database: an update. *Nucleic Acids Res.*, **42**, D1083–D1090.
- Bateman, A.M.M., O'Donovan, C., Magrane, M., Apweiler, R., Alpi, E., Antunes, R., Arganiska, J., Bely, B., Bingley, M., Bonilla, C. *et al.* (2015) UniProt: a hub for protein information. *Nucleic Acids Res.*, **43**, D204–D212.
- Gong, S., Worth, C., Cheng, T.K. and Blundell, T. (2011) Meet me halfway: when genomics meets structural bioinformatics. *J. Cardiovas. Transl. Res.*, **4**, 281–303.
- Hanson, R.M. (2010) Jmol – a paradigm shift in crystallographic visualization. *J. Appl. Cryst.*, **43**, 1250–1260.
- Tanramluk, D., Schreyer, A., Pitt, W.R. and Blundell, T.L. (2009) On the origins of enzyme inhibitor selectivity and promiscuity: a case

- study of protein kinase binding to staurosporine. *Chem. Biol. Drug Des.*, **74**, 16–24.
23. Pulkoski-Gross,A., Li,J., Zheng,C., Li,Y., Ouyang,N., Rigas,B., Zucker,S. and Cao,J. (2015) Repurposing the antipsychotic trifluoperazine as an antimetastasis agent. *Mol. Pharmacol.*, **87**, 501–512.
 24. Yeh,C.-T., Wu,A.T.H., Chang,P.M.H., Chen,K.-Y., Yang,C.-N., Yang,S.-C., Ho,C.-C., Chen,C.-C., Kuo,Y.-L., Lee,P.-Y. *et al.* (2012) Trifluoperazine, an antipsychotic agent, inhibits cancer stem cell growth and overcomes drug resistance of lung cancer. *Am. J. Respir. Crit. Care Med.*, **186**, 1180–1188.
 25. Laskowski,R.A., Hutchinson,E.G., Michie,A.D., Wallace,A.C., Jones,M.L. and Thornton,J.M. (1997) PDBsum: a web-based database of summaries and analyses of all PDB structures. *Trends Biochem. Sci.*, **22**, 488–490.
 26. Ihlenfeldt,W., Takahashi,Y., Abe,H. and Sasaki,S. (1994) Computation and management of chemical properties in CACTVS: An extensible networked approach toward modularity and compatibility. *J. Chem. Inf. Comput. Sci.*, **34**, 109–116.
 27. Ihlenfeldt,W.D., Voigt,J.H., Bienfait,B., Oellien,F. and Nicklaus,M.C. (2002) Enhanced CACTVS browser of the Open NCI Database. *J. Chem. Inf. Comput. Sci.*, **42**, 46–57.